Update of stock structure of humpback whales in the Antarctic feeding grounds as revealed by microsatellite DNA data (SC/65b/J31Rev)

Naohisa Kanda, Mutsuo Goto, and Luis A. Pastene

The Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo 104-0055, Japan

Contact email: kanda@cetacean.jp

ABSTRACT

This paper is a revised version of SC/F14/J31 presented to the JARPAII Special Permit Review Workshop, which takes the workshop's recommendations into consideration.

A total of 581 humpback whale biopsy samples obtained from Areas III to VI during surveys of the JARPA and JARPAII up to 2010/11 season as well as International Decade for Cetacean Research/Southern Ocean Whale and Ecosystem Research (IDCR/SOWER) were analyzed using 14 microsatellite DNA loci in order to describe their stock structure in the Antarctic feeding ground. The number of the loci used was increased from six in the paper submitted to the JARPA Review workshop in 2006. In three of 37 cases of duplicate sampling, the second samples were collected at least a day apart (1 day, 9 years, and 11 years). Paternity analysis of 13 calf and mother pairs failed to detect any potential fathers in the samples for the calves. After exclusion of some of these samples, 528 were used for further analyses at stock level. These individuals were divided into four groups based on the IWC management areas: IIIE (N=93), IV (N=218), V (N=153) and VIW (N=64). Heterogeneity tests for temporal and special genetic differences were conducted for each of the three different cases: samples of females+males, females only, and males only. F_{ST} values for these tests were also calculated. Although a few cases of small temporal genetic differences were detected within the areas, major genetic differences were observed among the different areas in all of the cases. In addition to that, stronger differentiation was seen in females than in males although both sexes basically showed similar pattern for the spatial differentiation. These results corresponded to those of the previous studies. Despite the increase of the number of loci, the level of the stock differentiation (F_{ST} = 0.003) was still too low to conduct a clustering analysis at the individual level.

With substantial increases in the numbers of the analyzed microsatellite loci and the biopsy samples, our genetic study again showed that humpback whales from the different stocks occupied the research areas with higher differentiation in females than in males. The level of the genetic differentiation among the areas was still so low that further analysis would require using samples from their breeding areas to better understand the stock structure in the feeding grounds. This study demonstrated one of the significant contributions of non-lethal part of the comprehensive large-scale JARPAII to acquire valuable information for effective management of large whales in the Antarctic.

KEYWORDS: ANTARCTIC, BIOPSY SAMPLING, FEEDING GROUNDS, GENETICS, HUMPBACK WHALE

INTRODUCTION

Humpback whales conduct seasonal migration between mid to high latitudinal waters in summer for feeding and low latitudinal waters in winter for breeding. In the Antarctic, humpback whales appear to congregate into five or six distinct feeding groups during the austral summer season that roughly match to IWC Management Areas I-VI (Mackintosh, 1965).

At the JARPA Review workshop in 2006, Pastene *et al*., (2006) conducted genetic analyses of humpback whales from the JARPA research area (Areas III to VI) using mitochondrial DNA and microsatellite DNA markers, and demonstrated that humpback whales from the different areas belonged to genetically different groups because of the observed genetic differences among them. In addition, it showed that the level of genetic

differentiation was larger in females than in males, strongly suggesting higher female phylopatry and large male dispersion in the feeding grounds. The next question to be addressed was whether or not these feeding groups were genetically distinct enough from each other that they could correspond to breeding stocks. Pastene *et al*. (2013) conducted mtDNA analysis using samples from both the feeding and breeding grounds, and showed the genetic differences observed among the areas were a consequence of the different proportions of the stocks from single stock occupancy in some areas to multi-stock occupancy in other areas.

This study used microsatellite DNA markers to analyze the same humpback whale samples from Areas III to VI used in Pastene *et al*. (2013). The main purpose of the study is to better understand their stock structure by increasing the number of the loci from Pastene *et al*. (2006). In that study, the use of the only six loci might have resulted in the observed low level of the genetic differentiation among the samples. The estimated F_{ST} value was about 0.005. At that level of differentiation, it was difficult to conduct clustering analysis at the individual base. In this paper, we increased the number of the microsatellite loci to 14. In addition, sample size was also substantially increased from the previous study because of the effort of JARPAII.

This paper is the revised version of SC/F14/J31 to cover the recommendations of the Panel of the Expert Workshop to review the JARPAII Special Permit Research Programme held at Tokyo from 24 - 28 of February, 2014. The panel recommends that SC/F14/J31 should provide locus-specific HW results with F_{IS} in Table 3 and F_{ST} per comparison in Tables 4-9.

MATERIALS AND METHODS

Samples

Table 1 shows the number of humpback whale biopsy samples used in this study separated by IWC management areas. A total of 581 samples were obtained from the JARPA and JARPAII surveys up to 2010/11 season as well as International Decade for Cetacean Research/Southern Ocean Whale and Ecosystem Research (IDCR/SOWER).

DNA extraction

In regard to our DNA data quality control under the IWC guidelines, see Kanda *et al*. (2014). Total DNA was extracted from 0.05 g of biopsy skin tissue using either the protocol of Sambrook *et al*. (1989) or GENTRA PUREGENE DNA extraction kit (QIAGEN). Extracted DNA was stored in the TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

Microsatellite analysis

Genetic variation at microsatellite loci was analyzed using 14 sets of primers: AC137, CA234 (Bérubé *et al*., 2005), EV1, EV14, EV37 (Valsecchi and Amos, 1996), GT23, GT195, GT271, GT310 (Bérubé *et al*., 2000), GATA28, GATA53, GATA98, GATA417, TAA31 (Palsbøll *et al*., 1997). All except EV1 and EV14 were designed specifically from humpback whales. Amplified products were run on a 6% polyacrylamide denaturating gel using a BaseStation 100 DNA fragment analyzer (Bio-Rad), and then sizes of visualized alleles were determined manually in relation to the internal size standard (Genescan 400HD, Life technologies) as well as the humpback whale's microsatellites of known size that were rerun on each gel. The sex of the whales was determined through co-amplification of SRY locus located on the Y chromosome and GT23, which is a slight modification from Abe *et al*. (2001). With this combination of loci, males show amplified products of both SRY and GT23 loci, while females show only GT23.

Data analysis

MICRO-CHECKER (van Oosterhout *et al*., 2004) was used to check for null alleles and reading/typing errors. The number of alleles per locus, allelic richness, and expected heterozygosity per locus was calculated using FSTAT 2.9.3 (Goudet, 1995). Statistical tests for the deviations from expected Hardy-Weinberg genotypic proportions for each of the loci were conducted, as well as their F_{IS} values (Weir and Cockerham, 1984) were calculated, using GENEPOP 4.0 (Rousset, 2008). Paternity analysis for assignment of mother and calf pairs to their potential fathers using genetic markers was conducted using CERVUS (Marshall *et al*., 1998). In order to examine genetic differences among samples, conventional hypothesis testing procedure was conducted using heterogeneity test in microsatellite allele frequencies among samples. A probability test (or Fisher's exact test) implemented in GENEPOP 4.0 (Rousset, 2008) was used to conduct the heterogeneity tests. Our null hypothesis to be tested is whether or not the samples came from a genetically same group of humpback whales. If statistically significant allele frequency differences exist, it could indicate these samples came from genetically different stocks of humpback whales. Statistical significance was determined using the chi-square value obtained from summing the negative logarithm of p-values over the 14 microsatellite loci (Sokal & Rohlf 1995). F_{ST} value was calculated using either GENEPOP 4.0 or FSTAT 2.9.3.

RESULTS AND DISCUSSION

Among 37 cases of the duplicate sampling, the second samples of three cases were collected at least a day apart (1 day, 9 years, and 11 years; Table 2). One case was first collected in Area IV and was recollected in Area III, while the other two cases were collected and re-collected in the same area. Because Pastene *et al*. (2013) showed the migration of whales from the same stock to Areas III and IV, our results were consistent to the feeding site fidelity of humpback whales in the Antarctic. Thirteen calf and mother pairs were used for paternity analysis, but no potential father was detected in the samples.

After exclusion of the second individuals from the duplicated samples and calves sampled with mothers, a total of 528 samples were used for further analyses. These individuals were divided into four groups based on the management areas: IIIE (N=93), IV (N=218), V (N=153) and VIW (N=64). The levels of genetic diversity of the area samples were represented as an average number of alleles per locus, average allelic richness, and expected average heterozygosity (Table 3). These values were similar among them as well as to those based on the six loci previously reported in Pastene *et al*. (2006). No evidence of statistical significant p-values were obtained from the tests for deviation from the Hardy-Weinberg expected genotypic proportions were observed at each of the loci after correction for multiple tests (Rice, 1989) as well as at all the loci combined (Table 3). F_{IS} values appeared to have no trends toward either homozygote excess or heterozygote excess (Table 3).

Heterogeneity tests for genetic differences among the samples from the different areas were conducted for three different cases: samples of females+males, females only, and males only, respectively (Tables 4-9). F_{ST} values for each of the comparisons were also calculated. For the females+males case, first, genetic differences among the different year samples within the areas were examined (Table 4). No evidence of genetic differences was found among the samples from the different years in Areas IIIE, IV, and V, while small level of genetic differences was observed in Area VIW. Because the observed significance in Area VI was small and none of the pair-wise comparisons among the year samples was significant (results not shown), the samples from the different years within each of the areas were combined for further analysis. F_{ST} values were indeed low ranging from -0.0007 to 0.0026. Evidence of the genetic differences was then detected among the samples from the different areas (Table 5). Pair-wise comparisons showed evidence of statistically significant differences for all possible area-pairs (Table 5). Despite the significant differences from the heterogeneity tests, F_{ST} values were low for among the all areas (0.0026) as well as for the pair-wise comparisons (0.0004 - 0.0037) (Table 5).

The heterogeneity tests for the females only and males only cases in general showed similar results to each other with clearer spatial structuring in females than in males (Tables 6-9). A yearly genetic difference was detected at Area V in females only case and at area VIW in males only case (Tables 6 and 8). In the females only case, two (94/95x98/99 and 98/99x08/09) out of the 21 pair-wise comparisons were significant with close-to 5% p-values (0.035 and 0.049). In the males only case, one (96/97x00/01) out of three pair-wise comparisons was significant with a 0.017 p-value. Because it was difficult to conclude whether or not these yearly differences had any biological meanings, we combined the year samples within the areas for the next tests. Evidence of the genetic difference among the different areas was detected in both cases. In the females only case, all of the pair-wise comparisons showed the statistically significant difference, whereas in the males only case, all except Areas IIIxV and Areas VxVI comparisons showed the statistically significant difference (Tables 7 and 9). Again, despite the significant differences from the heterogeneity tests, F_{ST} values were low for among the all areas (0.0025 for the females only and 0.0028 for males only) as well as for the pair-wise comparisons (-0.0002 - 0.0042; 0.0009 - 0.0035).

The results of this study were consistent to those previously reported based on six microsatellite loci as whale groups occupying the different areas were genetically different from each other with stronger differentiation in females than in males. Pastene *et al.* (2013; see also 2006) analyzed mitochondrial DNA variations on humpback whales from both the feeding and breeding grounds to better describe their stock structure, and showed that IV from 80˚E to 120˚E was occupied by one stock (Western Australia stock) and V from 140˚E to 160˚E by another stock (Eastern Australia stock). The rest of the areas were mixing areas of the adjacent stocks. Despite the substantial increase in the number of the loci, the level of the stock differentiation (F_{ST} around 0.003) was still too low to conduct a clustering analysis at the individual level. In such a situation, it was difficult for this study to further distinguish between stock mixing and stock core areas without using the data from the breeding areas. Future microsatellite study should use the samples from the breeding areas. Nevertheless, substantial increases in the numbers of the analyzed microsatellite loci and the biopsy samples allowed us to confirm our previous conclusion on the humpback whale stock structure in the Antarctic. This study demonstrated one of the significant contributions of non-lethal part of the comprehensive large-scale JARPAII to acquire valuable information for effective management of large whales in the Antarctic.

ACKNOWLEDGEMENTS

We would like to thank all the researchers and crew members of JARPA/JARPA II and IDCR/SOWER surveys involved in the sampling of the Antarctic biopsy samples used in this study. We also would like to thank H. Hatanaka for his useful comments on an earlier version of this manuscript and the Panel member of the review workshop for their constructive comments and advices for the manuscript.

REFERENCES

Abe, H., Goto, M., Pastene, L.A., Dewa, K., Naito, E. 2001. Practical use of multiplex fluorescent PCR for cetacean sex identification. *Mar. Mammal Sci.* 17: 657-664.

Bérubé, M., Jørgensen, H., Mcewing, R., and Palsbøll, P.J. 2000. Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeanglliae*. *Mol. Ecol.* 9:2181-2183.

Bérubé, M., Rew, M.B., Skaug, H., Jørgensen, H., Robbins, J., Best, P., Sears, R., and Palsbøll, P.J. 2005. Polymorphic microsatellite loci isolated from humpback whale, *Megaptera novaeanglliae* and fin whale, *balaenoptera physalus*. *Cons. Genet*. 6: 631-636.

Buchanan, F.C., Friesen, M.K., Littlejohn, R.P., and Clayton, J.A. 1996. Microsatellites from beluga whale *Delphinapterus leucas*. *Mol. Ecol.* 5:571-575.

Goudet, J. 1995. FSTAT, version 1.2: a computer program to calculate F-statistics. *J. Hered*. 86:485-486.

Kanda, N., Goto, Oikawa, H., and Pastene, L.A. 2014. A note on sampling and laboratory procedure protocols of the genetic work at the Institute of Cetacean Research. Paper SC/F14/J27 presented to the JARPAII Review Workshop, Tokyo, February 2014.

Mackintosh, N.A. 1965. *The stocks of whales*. Fishing News (Books) Ltd., London. 232pp.

Marshall, T.C., Slate, J., Kruuk, L., and Pemberton, J.M. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7: 639-655.

Pastene, L.A., Goto, M., Nishiwaki, S., Yoshida, H. and Kanda, N. 2006. Genetic characteristics and population structure of humpback whales in the Antarctic feeding ground as revealed by mitochondrial DNA control region seqwuencing and microsatellite analyses. Paper SC/D06/J31 presented to the JARPA Review Workshop, December 2006 (unpublished). 14pp.

Pastene, L.A., Kitakado, T., Goto, M., and Kanda, N. 2013. Mixing rates of humpback whales of Stocks D, E and F in the Antarctic feeding grounds based on mitochondrial DNA analyses. Paper SC/65a/SH13 presented to the IWC Scientific Committee, June 2013, Cheju Island, Korea (unpublished).

Palsbøll, P.J., Bérubé, M., Larsen, A.H., and Jørgensen, H. 1997. Primers for the amplification of tri-and tetramer microsatellite loci in baleen whales. *Mol. Ecol.* 6:893-895.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.

Rousset, F. 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resources* 8:103-106.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning: A laboratory manual*. 2nd edn. Cold Spring Harbor Laboratory, New York, USA.

Sokal, R.R., and Rohlf, F.J. 1995. *Biometry: the Principles of Statistics in Biological Research*. Freeman and Company, New York.

Valsecchi, E., and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* 5:151-156.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P., and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4:535-538.

Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Table 1. The number of humpback whale biopsy samples usid in this study by surveys and areas.

Survey	Survey area			
	IIIE	IV	V	VIW
JARPA	56 (33, 23)	132 (58, 74)	106 (56, 50)	46 (18, 28)
JARPAII	14 (6, 8)	49 (26, 23)	61 (36, 25)	3 (0, 3)
IDCR/SOWER	36 (24, 12)	51 (30, 21)	7 (3, 4)	20 (12, 8)
All	106 (63, 43)	232 (114, 118)	174 (95, 79)	69 (30, 39)

Total (female, male)

Table 2. Sampling dates and lcoations of the matched samples.

Sex	Sampling	lat.	long.	Cap.→recap.
Female	1994/2/12	66.54	70.07E	IV → IIIE, 9 yrs later
	2003/12/14	60.31S	48.57E	
Female	1997/1/25	61.44S	151.42E	V → V, 11 yrs later
	2008/12/13	63.05S	138.26E	
Male	2002/2/12	64.08S	96.58E	IV → IV, next day
	2002/2/13	64.06S	97.24E	

Table 3. Genetic diversity indices estimated from all samples within each of survey areas.

Genetic indice	Survey area							
	IIIE		IV		V		VIW	
No. alleles per locus	10.6		11.9		11.9		10.1	
Allelic richness per locus	10.2		10.3		10.7		10.1	
Heterozygosity	0.758		0.752		0.755		0.746	
Hardy-Weinberg / Fis								
GATA417	0.960	-0.016	0.618	0.002	0.063	0.026	0.044	0.094
GATA28	0.889	-0.046	0.909	-0.025	0.749	-0.004	0.666	-0.038
GATA98	0.023	0.081	0.051	0.016	0.403	-0.037	0.438	0.011
TAA31	0.459	0.030	0.910	0.041	0.881	0.028	0.331	-0.038
GATA53	0.512	-0.030	0.432	-0.009	0.644	-0.063	0.981	-0.049
GT23	0.376	0.025	0.223	0.007	0.797	0.039	0.751	-0.018
GT195	0.725	0.000	0.624	-0.006	0.384	0.039	0.344	0.099
AC137	0.683	-0.018	0.740	-0.005	0.605	-0.010	0.446	0.059
EV1	0.141	-0.099	0.392	0.039	0.401	0.006	0.115	-0.007
GT271	0.675	-0.033	0.773	0.022	0.620	-0.049	0.113	-0.022
CA234	0.231	-0.086	0.872	0.003	0.533	0.041	0.652	0.070
GT310	0.470	-0.075	0.065	0.038	0.618	-0.022	0.750	-0.067
EV14	0.740	0.012	0.057	0.018	0.473	0.004	0.276	-0.016
EV37	0.904	-0.022	0.638	-0.018	0.070	0.021	0.024	-0.006
All	0.677		0.451		0.668		0.157	

Table 4. Results (p-values) of heterogeneity tests for year differences within areas and Fst: Females+Males.

	IIIE	IV	V	VIW
p-value	0.814	0.555	0.204	**0.049**
Fst	-0.0052	-0.0007	0.0015	0.0026

Bold p-values indicate statistically significance.

Table 5. Results (p-values) of heterogeneity tests among areas and Fst: Females+Males.

	Among areas	IIIExIV	IIIExV	IVxV	IIIxVIW	IVxVIW	VxVIW
p-value	**Highly sig.**	**>0.001**	**>0.001**	**>0.001**	**>0.001**	**>0.001**	**0.0471**
Fst	0.0026	0.0033	0.0037	0.0029	0.0012	0.0018	0.0004

Bold p-values indicate statistically significance.

Table 6. Results (p-values) of heterogeneity tests for year differences within areas and Fst: Females only.

	IIIE	IV	V	VIW
p-value	0.595	0.510	**0.015**	xxx
Fst	-0.002	0.000	0.008	xxx

Bold p-values indicate statistically significance. XXX indicates no comparison due to the small sample sizes.

Table 7. Results (p-values) of heterogeneity tests among areas and Fst: Females only.

	Among areas	IIIExIV	IIIExV	IVxV	IIIxVIW	IVxVIW	VxVIW
p-value	**>0.001**	**0.014**	**>0.001**	**>0.001**	**0.019**	**0.046**	**0.045**
Fst	0.0025	0.0035	0.0042	0.0027	-0.0002	-0.0002	0.0002

Bold p-values indicate statistically significance.

Table 8. Results (p-values) of heterogeneity tests for year differences within areas and Fst: Males only.

	IIIE	IV	V	VIW
p-value	0.414	0.641	0.284	**0.020**
Fst	-0.0037	-0.0018	0.0030	0.0085

Bold p-values indicate statistically significance.

Table 9. Results (p-values) of heterogeneity tests among areas and Fst: Males only.

	Among areas	IIIExIV	IIIExV	IVxV	IIIxVIW	IVxVIW	VxVIW
p-value	**>0.001**	**0.009**	0.067	**0.007**	**0.022**	**0.001**	0.163
Fst	0.0028	0.0035	0.0019	0.0032	0.0016	0.0035	0.0009

Bold p-values indicate statistically significance.